JOURNAL OF MULTIVARIATE ANALYSIS 22, 212–229 (1987)

Approximations for Multivariate U-Statistics

F. GÖTZE

Universität Bielefeld, Bielefeld, West Germany

Communicated by J. Pfanzagl

Edgeworth approximations for multivariate U-statistics hold up to the order $o(n^{-1/2})$ under moment conditions and the assumption that the projection of the U-statistic to sums of i.i.d. random vectors is strongly nonlattice. © 1987 Academic Press, Inc.

1. Introduction and Results

Let $X_1, X_2, \ldots, X_N$ be independent and identically distributed (i.i.d.) random variables assuming values in a measurable space $(\mathscr{X}, \mathscr{A})$ with common distribution P. Let $H: \mathscr{X} \times \mathscr{X} \to \mathbb{R}^k$, $k \geqslant 1$, integer, denote a vector-valued kernel $H(x, y) = (H_1(x, y), \ldots, H_k(x, y))$, which is symmetric in its arguments, i.e., $H(x, y) = H(y, x)$. For $N \geqslant 2$, a k-variate U-statistic of degree 2 is defined as

$$U_N \triangleq \binom{N}{2}^{-1} \sum_{\mu=1}^{N} \sum_{\nu=\mu+1}^{N} H(X_\mu, X_\nu). \tag{1.1}$$

We shall write $v \cdot w = v_1 w_1 + \cdots + v_k w_k$ for the euclidean scalar product of vectors $u, v \in \mathbb{R}^k$ and let $\|v\| \triangleq (v \cdot v)^{1/2}$ denote the euclidean norm. We assume throughout that

$$EH(X_1, X_2) = 0 \qquad \text{and} \qquad E\,\|H(X_1, X_2)\|^2 < \infty. \tag{1.2}$$

We define

$$g(X_1) \triangleq E(H(X_1, X_2) \mid X_1) \tag{1.3}$$

Received July 8, 1985; revised July 18, 1986.

AMS 1980 subject classification: Primary 60F05.

Key words and phrases: Berry–Esseen bound, Edgeworth expansion, moment condition, U-statistic.

0047-259X/87 $3.00

$$h(X_1, X_2) \triangleq H(X_1, X_2) - g(X_1) - g(X_2) \tag{1.4}$$

$$\hat{U}_N \triangleq 2N^{-1} \sum_{j=1}^{N} g(X_j)$$

$$R_N \triangleq \binom{N}{2}^{-1} \sum_{\mu=1}^{N} \sum_{\nu=\mu+1}^{N} h(X_\mu, X_\nu),$$

which leads to the representation

$$U_N = \hat{U}_N + R_N, \tag{1.5}$$

where the term R_N is asymptotically negligible compared to $\hat{U}_N$.

In order to obtain a nondegenerate limit distribution introduce

$$U_N^* \triangleq \tfrac{1}{2} N^{1/2} U_N. \tag{1.6}$$

Let V denote the covariance matrix of the vector $g(X_1)$ under P. Hoeffding [10] has shown that, as $N \to \infty$, the distribution of U_N^* converges weakly to the multivariate $N(0, V)$-normal distribution, provided (1.2) holds and V is positive definite.

The speed of convergence to normality for univariate U-statistics has been extensively studied in recent years. Berry–Esseen bounds $O(N^{-1/2})$ for the difference of distribution functions were established under conditions of increasing generality by Bickel [2], Chan and Wierman [7], Callaert and Janssen [4], and Helmers and van Zwet [9].

Better approximations for the distribution of U_N are obtained by means of Edgeworth expansions up to an error of order $o(N^{-1/2})$ or $o(N^{-1})$ which are relevant for comparisons of statistical procedures.

To be specific, let

$$D \triangleq \left(\frac{\partial}{\partial a_1}, \ldots, \frac{\partial}{\partial a_k}\right) \qquad \text{and} \qquad g(x) \cdot D \triangleq \sum_{j=1}^{k} g_j(x) \frac{\partial}{\partial a_j},$$

and define the cumulant differential operator

$$\kappa_3(D) \triangleq E(g(X_1) \cdot D)^3 + 3Eh(X_1, X_2)(g(X_1) \cdot D)(g(X_2) \cdot D) \tag{1.7}$$

and define the (possibly signed) measure Ψ_N approximating the distribution of U_N^*, by means of its Lebesgue-density ψ_N in $\mathbb{R}^k$,

$$\psi_N(a) \triangleq (1 + \tfrac{1}{6} N^{-1/2} \kappa_3(-D))\, \varphi_{0,V}(a), \qquad a \in \mathbb{R}^k, \tag{1.8}$$

where $\varphi_{0,V}(a)$ denotes the Lebesgue-density of $N(0, V)$.

Higher-order approximations up to an error of order $o(N^{-1})$ for univariate U-statistics were studied by Callaert *et al.* [5] and Bickel *et al.*

[3]. For multivariate U-statistics a Berry–Esseen bound has been proved by Carmichael [6] under moment assumptions on $\|H(x, y)\|$ of order $k+4$.

The aim of the present paper is to establish Berry–Esseen bounds and Edgeworth expansions up to the order $o(N^{-1/2})$ under less restrictive moment conditions.

In order to obtain uniform error bounds for the Edgeworth expansions we have to introduce some constants. Assume there exist constants $a>0$, $C>0$ and continuous nonincreasing functions χ_j: $[0, \infty) \to [0, \infty)$, $j=1, 2$, satisfying

$$\lim_{x \to \infty} \chi_1(x)=0 \tag{1.9}$$

$$0<\chi_2(x) \leqslant 1 \qquad \text{for every } x \geqslant a \tag{1.10}$$

$$E\,\|H(X_1, X_2)\|^3\, I(\|H(X_1, X_2)\| \geqslant x) \leqslant \chi_1(x) \qquad \text{for every } x>0 \tag{1.11}$$

the distribution of $g(X_1)$ is uniformly strongly nonlattice, i.e.,

$$|E \exp[it \cdot g(X_1)]| \leqslant 1-\chi_2(\|t\|), \qquad \text{for every} \quad \|t\| \geqslant a>0. \tag{1.12}$$

Let $L_{C,\nu}$ denote a system of Borel sets in $\mathbb{R}^k$ defined as

$$L_{C,\nu} \triangleq \{A \in B^k \colon \Phi_{0,V}((\partial A)^\varepsilon) \leqslant C\varepsilon+\nu, \quad \text{for every} \quad \varepsilon>0\}, \qquad \nu>0, \tag{1.13}$$

where $(\partial A)^\varepsilon=\{x \in \mathbb{R}^k \colon \exists\, y \in \partial A\ \|y-x\| \leqslant \varepsilon\}$. We shall prove

THEOREM 1.14. *Suppose that there exist a constant $C>0$ and functions $\chi_j, j=1, 2$, such that* (1.2) *and* (1.9)–(1.12) *are fulfilled. Then there exist a sequence $\varepsilon_N \downarrow 0$* and *a constant K depending on C, $\chi_j(\cdot), j=1, 2$, only such that for $N=1, 2, \ldots$, $\nu_N \triangleq \varepsilon_N N^{-1/2}$,*

$$\sup_{A \in L_{C,\nu_N}} |P(U_N^* \in A)-\Psi_N(A)| \leqslant K\varepsilon_N N^{-1/2}, \tag{1.15}$$

where U_N^* and Ψ_N are given by (1.6) and (1.7)–(1.8) and L_{C,ν_N} is defined in (1.13).

Remark. The assumption (1.12) that $g(X_1)$ has a strongly nonlattice distribution already entails that the limit distribution of U_N^* is uniformly nondegenerate, i.e., the smallest eigenvalue λ_V of the asymptotic covariance matrix V satisfies

$$\lambda_V \geqslant a^{-2} \inf_{\|t\|=1} (1-\chi_2(at)^2)>0.$$

This is an immediate consequence of

$$1-|E\exp[it\cdot g(X_1)]|^2=2\,E\sin^2[t\cdot(g(X_2)-g(\bar{X}_1))]\leqslant E(t\cdot g(X_1))^2,$$

where $\bar{X}_1$ denotes an independent copy of X_1.

As a corollary we have

THEOREM 1.16. *Assume that* (1.2) *holds and*

(i) $E\,\|H(X_1,X_2)\|^3<A<\infty$,

(ii) *the smallest eigenvalue of* $E\,g(X_1)^Tg(X_1)$ *is larger than* $\delta>0$.

Then there exist constants $d_1, d_2>0$ *depending on* δ, A, C *such that with* $\nu_N\triangleq d_1N^{-1/2}$

$$\sup_{A\in L_{C,\nu_N}}|P(U_N^*\in A)-\int_A\varphi_{0,V}(x)\,d^kx|\leqslant d_2N^{-1/2},$$

uniformly for every $H(X_1,X_2)$, *satisfying the conditions above.*

The conditions (1.9)–(1.12) allow for the dependence of $h(x,y)$ and P on N, which is desirable for many applications. Consider, for example, the statistic [11]

$$T_N\triangleq N^{-1/2}\sum_{j=1}^N(f_0(X_j)+N^{-1/2}f_1(X_j))+N^{-3/2}\sum_{1\leqslant j<m\leqslant N}f_2(X_j,X_m),$$

where $f_j, h=1,2,3$ denote $\mathbb{R}^k$-valued functions, such that $Ef_j(X_1)=0, j=1,2,$ $E(f_2(X_1,X_2)\,|\,X_2)=0$ a.s. and $f_2(x,y)$ is symmetric. Assume that f_2 satisfies (1.11) and condition (1.12) holds for f_0. If, furthermore,

$$E\,\|f_1(X_1)\|^{3/2}\leqslant C_1<\infty, \tag{1.17}$$

then we have

COROLLARY 1.18. *Under the same conditions as in Theorem* 1.14 *and condition* (1.17) *there exists a sequence* $\varepsilon_N\downarrow 0$ *and a constant* K *depending on* $\chi_j(\cdot)$, C *and* C_1 *only such that*

$$\sup_{A\in L_{C,\nu_N}}|P(T_N\in A)-\tilde{\Psi}_N(A)|\leqslant K\varepsilon_N N^{-1/2},\qquad N\geqslant 2,\ \nu_N\triangleq\varepsilon_N N^{-1/2}, \tag{1.19}$$

where L_{C,ν_N} *is defined as in* (1.13) *and*

$$\tilde{\Psi}_N(A)\triangleq\int_A\{1+\tfrac{1}{6}N^{-1/2}\kappa_3(-D)\}\,\varphi_{0,V_N}(x)\,d^kx,$$

$V_N \triangleq \operatorname{cov}(f_0 + N^{-1/2} f_1)$, and $\kappa_3(D)$ *is defined with respect to* $f_0(X)$ *and* $f_2(x, y)$ *as in* (1.7).

The nonuniform versions of conditions (1.9)–(1.12) are simply

$$E \, \|H(X_1, X_2)\|^3 < \infty \tag{1.20}$$

$$|E \exp[it \cdot g(X_1)]| < 1 \qquad \text{for every} \quad t \neq 0. \tag{1.21}$$

COROLLARY 1.22. *Assume that conditions* (1.2) *and* (1.20)–(1.21) *are satisfied. Then* (1.15) *holds.*

In the univariate case the moment condition (1.20) can be relaxed. It is sufficient to require an absolute moment of order $2 + \delta$, $\delta > 0$, for $h(X_1, X_2)$ (see Bickel *et al.* [3, Theorem 1.2]).

The proofs in the multidimensional case require different techniques, since there is no direct analogue to the Fourier inversion of differences of distribution functions in several dimensions.

Nevertheless, truncation techniques make it possible to reduce the estimation problem to comparing derivatives of characteristic functions. Here, the basic estimation techniques of Bickel *et al.* [3] are essentially applicable because of prior truncation.

On the other hand, the truncation of the kernel $h(x, y)$ introduces an error in the distribution function of the U-statistic which seems to require the assumption of a third moment of $H(x, y)$ for an error of size $o(N^{-1/2})$.

In this paper we confined ourselves to the approximation order $o(N^{-1/2})$. Assuming that $g(x)$ fulfills a multivariate Cramér condition and that

$$\dim_{\mathbb{R}} \left\langle (f_1, ..., f_k) \in L_2(\mathscr{X}, P)^k \colon \sum_{j,p=1}^{k} E f_j(X) \, h_j(X, Y) \, h_p(Y, Z) f_p(Z) > 0 \right\rangle$$

is sufficiently large, approximations up to $o(N^{-1})$ can be proved as in Bickel *et al.* [3]. Unfortunately the truncation techniques used in this paper do not allow one to prove the critical relation

$$E \left(\sum_{j=1}^{m} \sum_{i=j+1}^{N} h(X_j, X_i) \right)^r = O((mN)^{r/2}), \qquad 1 \leqslant m \leqslant N, \, 2 \leqslant r \leqslant k+1$$

for truncated h unless we assume that a large number ($r \geqslant k$) of moments exist. This estimate is an essential step used in several parts of the latter paper.

To prove Theorem 1.14 we shall reduce the problem to comparing c.f. In Section 3, for the reader's convenience we will state in a series of Lemmas some basic inequalities of Bhattacharya and Sweeting dealing with the multivariate case.

After introducing the necessary truncation techniques in Lemmas 2.5 and 2.7, we estimate the difference of the c.f. of U_N^* and Ψ_N for "small" frequencies in Lemma 3.5. In Lemma 3.10 these c.f. are shown to have negligible size for "large" frequencies up to $K_N \cdot N^{1/2}$. One technical result, the estimation of *arbitrary* moments of truncated U-statistics, which is used in Lemmas 3.5 and 3.10 and might be of wider interest, has been deferred to an appendix. In the following let c denote a generic positive constant depending on the constants introduced in (1.9)–(1.13).

2. Proofs

Reduction to Truncated Statistics

We shall use three types of modifications of the given statistic U_N^* of Theorem 1.14.

(1) Truncation of $h(X_j, X_l)$ at norm N.

(2) Truncation of $E^{1/3}(\|H(X_1, X_2)\|^3 \mid X_1)$ at $N^{1/2}$.

(3) Omission of the terms $N^{-1/2} \sum_{j=1}^{m} \sum_{l=j+1}^{N} h(X_j, X_l)/(N-1)$, $m = o(N^{1/2})$ in U_N^*.

Define

$$\begin{aligned} h''(x, y) &\triangleq h(x, y)\, I(\|h(x, y)\| \leqslant N) \\ M(x) &\triangleq E^{1/3}\, \|H(x, X_1)\|^3 \end{aligned} \tag{2.1}$$

$$\begin{aligned} B_N &\triangleq \{M(X_j) \leqslant N^{1/2}, j = 1, ..., N\} \\ E'f &\triangleq E(f \mid B_N), \qquad P'(A) \triangleq P(A \mid B_N) \end{aligned} \tag{2.2}$$

for bounded measurable f and A.

Put

$$\begin{aligned} g'(x) &\triangleq g(x) - E'g(X_1), \qquad j = 1, ..., N \\ h'(x, y) &\triangleq h''(x, y) - E'h''(x, X_1) - E'h''(X_1, y) + E'h''(X_1, X_2) \end{aligned} \tag{2.3}$$

and

$$T'_N \triangleq N^{-1/2} \sum_{j=1}^{N} g'(X_j) + N^{-1/2} \sum_{\mu=m}^{N} \sum_{\nu=\mu+1}^{N} h'(X_\mu, X_\nu)/(N-1),$$

where

$$m \triangleq [N^{1/2}\eta_N], \qquad \eta_N \triangleq (\log N^\kappa)^{-1}, \qquad \kappa > 0 \text{ to be determined later.}$$

Let Ψ'_N denote the measure with density (1.8), where all moments of g and h have been replaced by those of g' and h' under E'. We claim

LEMMA 2.5. $\sup_{A\in B^k} |\Psi_N(A)-\Psi'_N(A)| = o(N^{-1/2})$.

Proof. For every $s\in\mathbb{R}^k$, $\|s\|=1$ we have

$$|E'(s\cdot g(X_2))^2 - E(s\cdot g(X_1))^2| = o(N^{-1/2}).$$

By (2.2) and (2.3) we conclude $V=E'g^T(X_1)\,g(X_1)+o(N^{-1/2})$. Hence, when λ resp. λ' denotes the smallest eigenvalue of V resp. $V' \triangleq E'g'^T(X_1)\,g'(X_1)$, we have by means of (2.13)

$$\lambda' \geqslant \lambda - o(N^{-1/2}). \tag{2.6}$$

Furthermore, for every $s_1, s_2, s_3 \in \mathbb{R}^k$, $\|s_j\|=1$

$$E\prod_{j=1}^{3}(g(X_j)\cdot s_j) = E'\prod_{j=1}^{3}(g'(X_j)\cdot s_j)+o(1)$$

and

$$\begin{aligned}&E(g(X_1)\cdot s_1)(h(X_1,X_2)\cdot s_2)(g(X_2)\cdot s_3)\\&\quad = E'(g'(X_1)\cdot s_1)(h'(X_1,X_2)\cdot s_2)(g'(X_2)\cdot s_3)+o(1)\end{aligned}$$

after some elementary estimations. Hence

$$|\Psi_N(A)-\Psi'_N(A)| \leqslant o(N^{-1/2})\int_A (1+\|X\|^3)\exp\left[-\frac{1}{4}x^TVx\right]d(x)\,d^kx,$$

where

$$\begin{aligned}0\leqslant d(x) &\leqslant \exp\left[-\frac{1}{4}x^TVx\right]\left(1-\exp\left[-\frac{1}{2}x^T(V'-V)\,x\right]\right)\\ &\leqslant 1 \qquad \text{provided } N \text{ is sufficiently large.}\end{aligned}$$

This proves the lemma.

The error in replacing U_N^* by T'_N is estimated in

LEMMA 2.7. *Let δ_N be defined as in* (2.9). *Then*

$$\begin{aligned}|P(U_N^*\in A)-\Psi_N(A)| &\leqslant \max_{|\delta|=\delta_N}|P'(T'_N\in A^\delta)-\Psi'_N(A^\delta)|+o(N^{-1/2})\\ &\quad+\max_{|\delta|=\delta_N}|\Psi_N((\partial A)^\delta)|,\end{aligned}$$

where for $\delta>0$

$$A^{-\delta}\triangleq((A^c)^\delta)^c \text{ and } A^c\triangleq\mathbb{R}^k\backslash A.$$

Proof. Since $P(h''(X_1, X_2) \neq h(X_1, X_2)) = P(\|h(X_1, X_2)\| > N)$, we have by Chebyshev's inequality

$$P(h''(X_\mu, X_\nu) \neq h(X_\mu, X_\nu),\ 1 \leqslant \mu, \nu \leqslant N) \leqslant N^2 \cdot N^{-3} E\, \|h(X_1, X_2)\|^3 I(\|h(X_1, X_2)\| > N). \quad (2.8)$$

Furthermore,

$$P(B_N^c) \leqslant N^{-1/2}[E\, M(X_1)^3 I(M(X_1) > N^{1/2})] \triangleq \delta_N. \quad (2.9)$$

Let T''_N denote U^*_N with h replaced by h''. Hence (2.8)–(2.9) imply

$$\sup_{A \in B^k} |P(U^*_N \in A) - P'(T''_N \in A)| = o(N^{-1/2}). \quad (2.10)$$

Furthermore ($\varepsilon_N > 0$, $\varepsilon_N \downarrow 0$),

$$\begin{aligned} P'(\|T''_N - T'_N\| > \varepsilon_N N^{-1/2}) &\leqslant P'\left(\left\| N^{-3/2} \sum_{j=1}^{N} (N-1)\, g_h(X_j) \right\| > \frac{1}{3} \varepsilon_N N^{-1/2}\right) \\ &\quad + P'\left(\left\| N^{-3/2} \sum_{\mu=1}^{m} \sum_{\nu=\mu+1}^{N} h'(X_\mu, X_\nu) \right\| > \frac{1}{3} \varepsilon_N N^{-1/2}\right) \end{aligned}$$

where

$$g_h(x) \triangleq E'h''(X_1, x) - E'h''(X_1, X_2), \quad (2.11)$$

provided

$$\left\| N^{-1/2} \sum_{j=1}^{N} E'g(X_j) \right\| + \left\| N^{-3/2} \sum_{\mu=1}^{N} \sum_{\nu=\mu+1}^{N} E'h''(X_\mu, X_\nu) \right\| \leqslant \frac{1}{3} \varepsilon_N N^{-1/2}. \quad (2.12)$$

Notice that

$$\begin{aligned} c_N &\triangleq P(M(X_1) \leqslant N^{1/2}) = 1 - o(N^{-3/2}) \\ \|E'g(X_1)\| &= \|E\, g(X_1)\, I(\|M(X_1)\| \leqslant N^{1/2})/c_N\| \\ &= \|-E\, g(X_1)\, I(\|M(X_1)\| > N^{1/2})/c_N\| \\ &\leqslant N^{-1/2} \delta_N \quad (2.13) \\ \|E'h''(X_1, X_2)\| &\leqslant \|E'h(X_1, X_2)\| + \|E'h(X_1, X_2)\, I(\|h(X_1, X_2)\| > N)\| \\ &\leqslant c_N^{-2}\, \|E\, h(X_1, X_2)\, I(M(X_j) \leqslant N^{1/2}, j = 1, 2)\| + o(N^{-2}) \\ &\leqslant c_N^{-2}\, E\, \|h(X_1, X_2)\|\, M(X_1)^2\, I(M(X_1) > N^{1/2})\, N^{-1} \\ &\leqslant 2\delta_N N^{-1/2}. \quad (2.14) \end{aligned}$$

Choosing

$$\varepsilon_N \triangleq 9 \cdot \max(\delta_N N^{1/2}, (\log N^\kappa)^{-1/3}) \qquad \text{proves (2.12).} \tag{2.15}$$

Using Chebyshev's inequality applied to (2.11) together with (2.13),

$$\begin{aligned}
&P'(\|T_N'' - T_N'\| > \varepsilon_N N^{-1/2}) \\
&\qquad \leqslant 9\, \varepsilon_N^{-2} N E' \|g_h(X_1)\|^2 + 9\, \varepsilon_N^{-2} N mN N^{-3} E' \|h'(X_1, X_2)\|^2 \\
&\qquad \leqslant 9\, \varepsilon_N^{-2} N(N^{-2} E' \|E(h''(X_2, X_1) \\
&\qquad\quad \times M(X_2)^2 I(M(X_2) > N^{1/2}) \mid X_1)\|^2) \\
&\qquad\quad + 4\delta_N^2 N^{-1} + N^{1/2} \varepsilon_N^3 N^{-2} 4E \|h(X_1, X_2)\|^2) \\
&\qquad \leqslant c\, \varepsilon_N N^{-1/2}, \qquad N \text{ sufficiently large.}
\end{aligned} \tag{2.16}$$

Hence with $\delta = \varepsilon_N N^{-1/2}$,

$$\begin{aligned}
P'(T_N'' \in A) - \Psi_N'(A) \leqslant P'(T_N' \in A^\delta) - \Psi_N'(A^\delta) + \Psi_N'(A^\delta) - \Psi_N'(A) \\
+ P'(\|T_N' - T_N''\| > \delta)
\end{aligned}$$

yields, together with a similar lower bound and (2.16),

$$\begin{aligned}
|P'(T_N'' \in A) - \Psi_N(A)| \leqslant \max_{\delta, -\delta} |P'(T_N' \in A^\delta) - \Psi_N(A^\delta)| \\
+ \max_{\delta, -\delta} |\Psi_N(A^\delta) - \Psi_N(A)| + o(N^{-1/2}).
\end{aligned}$$

This together with (2.16) and Lemma 2.5 proves Lemma 2.7.

3. Characteristic Functions

The reduction of the estimation problem to comparing c.f. can be achieved using

LEMMA 3.1 (Bhattacharya and Sweeting). *Let $\hat{\chi}(t)$ denote the c.f. of the density*

$$\chi(x) = c_k \prod_{j=1}^{k} \frac{\sin(x_j/b)^{2p}}{x_j^{2p}},$$

$p \geqslant k$, $b > 0$ *chosen such that* supp $\hat{\chi} \subset [-1, 1]^k$. *Let* $a > 0$ *such that* $\alpha \triangleq \int_{\|x\| \leqslant a} \chi(x)\, d^k x > \frac{1}{2}$.

Then for any statistic T'_N and $A \in B^k$, there exist constants $c_j > 0$ depending on k, p, b, and a such that

$$|P(T'_N \in A) - \Psi_N(A)| \leqslant c_1 \max_{|\alpha| \leqslant k+1} \int |D^\alpha \{\hat{\chi}(t/t_N)(E \exp[it \cdot T'_N] - \hat{\Psi}_N(t))\}| \, d^k t + c_2(t_N^{-(p-1)} + (\alpha^{-1} - 1)^{r_N}) + c_3 N_{0,(1/2)V}((\partial A)^{2a/t_N}), \quad (3.2)$$

where $r_N = \lfloor t_N^{1/2} a^{-1} \rfloor$ and $t_N \uparrow \infty$, provided the covariance matrix V of the normal distribution N is nondegenerate.

Proof. The result is a combination of Lemma 5 of Sweeting [12, p. 37] and Lemma 11.6 of Bhattacharya and Ranga Rao [1, p. 98] together with some elementary estimations like

$$\Psi_N(A) \leqslant c_4 \int_A (1 + \|x\|^3) \, \varphi_{0,(1/2)V}(x) \, \varphi_{0,(1/2)V}(x) \, d^k x$$
$$\leqslant c_5 N_{0,(1/2)V}(A), \qquad c_4, c_5 > 0.$$

In order to estimate the derivatives of c.f. of sums of i.i.d. random vectors $g'(X_j)$, we may use results of Bhattacharya and Ranga Rao [1].

Let $\gamma_N(t) \triangleq E' \exp[it \cdot g'(X_1) N^{-1/2}]$.

LEMMA 3.3. *Assume $\|g'(X_1)\| \leqslant N^{1/2}$ a.s. and conditions (1.11) and $E'g'(X_1) = 0$. Then there exists $C, \delta > 0$ such that*

(i) $|D^\alpha \gamma_N(t)^N| \leqslant c(1 + \|t\|^{|\alpha|}) \exp[-\delta \|t\|^2]$

provided N is sufficiently large;

(ii) $$|D^\alpha(\gamma_N(t)^N - \exp[-\tfrac{1}{2} t^T V' t]\{1 + \tfrac{1}{6} N^{-1/2} E'(g'(X_1) \cdot t)^3\})| \leqslant o(N^{-1/2})[\|t\|^{(3-|\alpha|)_+} + \|t\|^{3+|\alpha|}] \exp[-\delta \|t\|^2]$$

for N sufficiently large, where V' denotes the covariance matrix of $g'(X_1)$ under E' and $(a)_+ \triangleq \max(a, 0)$.

Choosing $t_N = N^{1/2} K_N$, $K_N \uparrow \infty$, in Lemma 3.1 we have

$$\operatorname{supp} D^\beta \hat{\chi}(t t_N^{-1}) \subset [-t_N, t_N]^k \qquad \text{and} \qquad |D^\beta \hat{\chi}(t t_N^{-1})| \leqslant c_\beta.$$

Put

$$\phi_N(t) \triangleq E' \exp[it \cdot T'_N] \qquad \text{with } T'_N \text{ as defined in (2.4).}$$

Then Lemma 3.1 and condition (1.14) together imply

$$\sup_{A \in L_{C,v_N}} |P(U_N^* \in A) - \Psi_N(A)|$$
$$\leqslant c \max_{|\alpha| \leqslant k+1} \int_{\{\|t\| < K_N N^{1/2}\}} |D^\alpha(\phi_N(t) - \hat{\Psi}'_N(t))| \, d^k t + o(N^{-1/2}), \tag{3.4}$$

where c depends on k, p, b, and a only.

The estimation of the c.f. $\phi_N(t) \triangleq E' \exp[it \cdot T'_N]$ is done in a series of lemmas.

Let $\varepsilon = 1/[2(k+4)]$. Then

LEMMA 3.5.

$$\int_{\|t\| < N^\varepsilon} |D^\alpha(\phi_N(t) - \hat{\Psi}'_N(t))| \, d^k t = o(N^{-1/2}).$$

Proof. The following arguments are relatively standard (compare Bickel *et al.* [3]). Let $\gamma_N(t) \triangleq E' \exp[it \cdot g'(X_1) N^{-1/2}]$. Then we may rewrite

$$\phi_N(t) = \gamma_N^N(t) + \gamma_N^{N-2}(t)\tbinom{N}{2} N^{-3/2} E' \exp[it \cdot (g'(X_1) + g'(X_2)) N^{-1/2}]$$
$$\times it \cdot h'(X_1, X_2) + R_1(t), \tag{3.6}$$

where

$$R_1(t) \triangleq -E'(\Delta_N \cdot t)^2 \exp[it \cdot S_N + it \cdot \Delta_N \eta]$$
$$S_N \triangleq N^{-1/2} \sum_{j=1}^{N} g'(X_j),$$
$$\Delta_N \triangleq N^{-1/2}(N-1)^{-1} \sum_{m \leqslant \mu < v}^{N} h'(X_\mu, X_v)$$

and η is uniformly distributed over [0, 1], independent of all other random variables.

Furthermore,

$$E' \exp[it \cdot (g'(X_1) + g'(X_2)) N^{-1/2}] \, h'(X_1, X_2) \cdot t$$
$$= E' \Bigg[\prod_{j=1}^{2} \{\exp[N^{-1/2} it \cdot g'(X_j)] - 1 - N^{-1/2} it \cdot g'(X_j)\}$$
$$+ 2N^{-1/2} it \cdot g'(X_2) \{\exp[N^{-1/2} it \cdot g'(X_1)] - 1 - N^{-1/2} it \cdot g'(X_1)\}$$
$$- N^{-1}(t \cdot g'(X_1))(t \cdot g'(X_2)) \Bigg] h'(X_1, X_2) \cdot t$$
$$\triangleq -N^{-1} E(t \cdot g'(X_1))(t \cdot g'(X_2))(t \cdot h'(X_1, X_2)) + R_2(t). \tag{3.7}$$

Using the multivariate expansion of c.f. of S_N, and Lemma 3.3(ii) together with (3.6), (3.7), and (2.6), we have for some $\delta > 0$

$$\begin{aligned}|D^\alpha[\phi_N(t) - \hat{\Psi}'_N(t)]| &\leqslant c\,\delta_N \exp[-\delta\,\|t\|^2](1+\|t\|^5)\\ &\quad + N^{-1/2}\,|D^\alpha\{(\gamma_N^{N-2}(t) - \exp[\tfrac{1}{2}t\cdot\Sigma'\cdot t])\\ &\quad \times E\,t\cdot g'(X_1)\,t\cdot g'(X_2)\,h'(X_1,X_2)\cdot t\}|\\ &\quad + |D^\alpha(N^{1/2}R_2(t)\cdot\gamma_N^{N-2}(t))| + |D^\alpha R_1(t)|\\ &\triangleq c\,\delta_N\exp[-\delta\,\|t\|^2](1+\|t\|^5) + I_1 + I_2 + I_3. \end{aligned} \tag{3.8}$$

By Lemma 3.3(ii) we immediately conclude $\int_{\|t\| < N^\varepsilon} I_1\, d^k t \leqslant c\,\delta_N$.

Estimation of I_2.

$$I_2 \leqslant \sum_{\beta+\gamma=\alpha} |D^\beta \gamma_N^{N-2}(t)|\,|D^\gamma R_2(t)|\,N^{1/2}, \qquad \text{where} \quad \beta,\gamma \geqslant 0$$

denote k-tuples of nonnegative integral numbers.

$$\begin{aligned}|D^\gamma R_2(t)| &\leqslant c \sup_{|\alpha| \leqslant k+1} |D^\alpha R_2(t)|\\ &\leqslant c \sup_\beta E' \prod_{j=1}^{2} \|g'(X_j)\,N^{-1/2}\|^{\beta_j}\,\|h'(X_1,X_2)\|(1+\|t\|^5)\\ &\leqslant c(1+\|t\|^5)(1+E\,\|g'(X_1)\|\,\|h'(X_1,X_2)\|\,\|g'(X_2)\|^{9/5})\,N^{-14/10},\\ &\qquad\qquad N \text{ sufficiently large.}\end{aligned}$$

Since $|N^{-1/2}g'(X_1)| \leqslant 2$ a.e., Hölder's inequality shows that

$$I_2 \leqslant c\exp(-\delta\,\|t\|^2)(1+\|t\|^5)\,N^{-9/10}. \tag{3.9}$$

Estimation of I_3.

$$\begin{aligned} I_3 = |D^\alpha R_1(t)| &\leqslant c(1+\|t\|^2)\,E\,\|\Delta_N\|^2(1+\|\Delta_N\|^{k+1})\,\|S_N\|^{k+1}\\ &\leqslant c(1+\|t\|^2)(N^{-1} + E^{(1+\lambda)^{-1}}\,\|\Delta_N\|^{(k+3)(1+\lambda)}\\ &\quad \times E^{1/K}\,\|S_N\|^{(k+1)K})\\ &\leqslant c(1+\|t\|^2)\,(N^{-(k+3)/2} + N^{-1})\end{aligned}$$

by Hölder's inequality with $\lambda < \frac{1}{5}$ and appropriately chosen K together with Lemma A.1 and Lemma 3.3(i) with $t = 0$. Hence,

$$\sup_{|\alpha|\leqslant k+1}\int_{\|t\|<N^{\varepsilon}}|D^{\alpha}(\phi_N(t)-\hat{\Psi}'_N(t))|\,d^kt$$

$$\leqslant o(N^{-1/2})+c\int_0^{N^{\varepsilon}}r^{k-1}(1+r^2)\,dr\,N^{-1}$$

$$\leqslant o(N^{-1/2})+c\,N^{\varepsilon(k+3)-1},\qquad N \text{ sufficiently large,}$$

$$=o(N^{-1/2}),\qquad \text{since } \varepsilon<1/[2(k+3)].$$

LEMMA 3.10. *For every integral nonnegative vector* α, $|\alpha|\leqslant k+1$, *we have*

$$\text{(i)}\quad \int_{\|t\|>N^{\varepsilon}}|D^{\alpha}\hat{\Psi}'_N(t)|\,d^kt=o(N^{-1/2})$$

$$\text{(ii)}\quad \int_{A_{N1}\cup A_{N2}\cup A_{N3}}|D^{\alpha}\phi_N(t)|\,d^kt=o(N^{-1/2}),$$

where $A_{N1}\triangleq\{t\colon N^{\varepsilon}<\|t\|\leqslant N^{1/4}\eta_N^{-1}\}$, $A_{N2}\triangleq\{t\colon N^{1/4}\eta_N^{-1}<\|t\|\leqslant\varepsilon N^{1/2}\}$, *and* $A_{N3}\triangleq\{t\colon\varepsilon N^{1/2}<\|t\|\leqslant K_N N^{1/2}\}$ *for some* $K_N\uparrow\infty$.

Proof. Part (i) is obvious since the covariance matrix of $g'(X_1)$ is nondegenerate for sufficiently large N by relation (2.6).

(ii) *First step* (the region $N^{\varepsilon}<\|t\|\leqslant N^{1/4}\eta_N^{-1}$). We expand $\phi_N(t)$ in terms of Δ_N as in Lemma 3.5.

$$D^{\alpha}\phi_N(t)=D^{\alpha}\left\{E'\left(\exp[it\cdot S_N]\sum_{j=0}^{r}(i\Delta_N\cdot t)^j/j!\right)+R_1(t)\right\}$$

$$|D^{\alpha}R_1(t)|\leqslant cE'\,\|\Delta_N\|^{r+1}(1+\|S_N\|^{k+1}+\|\Delta_N\|^{k+1})(1+\|t\|^{r+1}).\tag{3.11}$$

By Lemma A.1 we have for some constant $K>0$

$$\int|D^{\alpha}R_1(t)|\,d^kt\leqslant K\int_{A_{N1}}N^{-(r+1)/2}(1+\|t\|^{r+1})\,d^kt\leqslant K\,N^{(k-r-1)/4}\eta_N^{-(k+r+1)}$$

$$=o(N^{-1/2})\tag{3.12}$$

choosing

$$r\triangleq k+1.\tag{3.13}$$

Furthermore,

$$|D^{\alpha}(\phi_N(t)-R_1(t))|\quad\max_{|\beta+\gamma|\leqslant k+1}\left\{E'\,|D^{\beta}\exp[it\cdot S_N]|\sum_{j=0}^{r}|D^{\gamma}(\Delta_N\cdot t)^j|\right\},$$

where β,γ denote nonnegative integral numbers.

Since Δ_N consists of less than N^2 summands,

$$|D^\alpha \phi_N(t)| \leqslant c \max_{|\beta+\gamma| \leqslant k+1} \sum_{j=0}^{r} \binom{N}{|\beta|} \left| E' D^\beta \exp[it \cdot S_{k+1+r}] \prod_{j,p}{}^{*} h'(X_j, X_p) \right|$$
$$\times \|t\|^{(j-|\gamma|)_+} N^{2j} |\gamma_N^{N-k-1-r}(t)| + |D^\alpha R_1(t)|,$$

where $\prod_{j,p}^*$ denotes a product of j factors with $1 \leqslant j \leqslant m$ and $1 \leqslant p \leqslant N$. By Lemmas 3.3(i) and A.1 and (2.1) as well as (2.2), we have

$$\leqslant c\, N^{|\alpha|}(1 + \|t\|^r N^{2r} \cdot N^r)\, |\gamma_N(t)|^{N-k-1-r} + |D^\alpha R_1(t)|.$$

Since

$$|\gamma_N(t)| \leqslant \exp(-c\, \|t\|^2/N), \qquad \text{for some} \quad c > 0,\ \|t\| < \varepsilon N^{1/2},$$

by Lemma 3.3(i), we conclude

$$\int_{A_{N1}} |D^\alpha \phi_N(t)|\, d^k t \leqslant N^{4(k+1)} \int_{A_{N1}} \|t\|^{k+1} \exp(-c\, \|t\|^2)\, d^k t + o(N^{-1/2})$$
$$= o(N^{-1/2}) \qquad \text{by the choice of } A_{N1}. \tag{3.14}$$

Second step (the region $N^{1/4}\eta_N^{-1} < \|t\| \leqslant \varepsilon N^{1/2}$). Since Δ_N depends on $X_m, ..., X_N$ only, where $m = [\eta_N N^{1/2}]$, we may write $L_N \triangleq \Delta_N + \sum_{j=m+1}^{N} N^{-1/2} g'(X_j)$ and

$$|D^\alpha \phi_N(t)| = |D^\alpha(\gamma_N^m(t)\, E' \exp[it \cdot L_N])|$$

$$\leqslant \sup_{\beta+\gamma=\alpha} |D^\beta \gamma_N^m(t)\, |E'\, \|L_N\|^{|\gamma|}, \qquad \beta, \gamma \geqslant 0 \text{ integral vectors}$$

$$\leqslant c \sup_{\beta \leqslant \alpha} m^{|\beta|}\, |\gamma_N(t)|^{m-|\beta|}, \qquad \text{using Lemma A.1}$$

$$\leqslant c\, N^{|\alpha|/2} \exp[-\lambda\, \|t\|^2 (m - |\alpha|)/N], \qquad \lambda > 0$$

$$\leqslant c\, N^{|\alpha|/2} \exp[-\lambda/2 \cdot \log(N^\kappa)] = o(N^{-(k+1)/2}) \qquad ((3.15)$$

if κ is chosen s.th. $\kappa < 2((k+1)/\lambda)$ and N is sufficiently large. Hence

$$\int_{A_{N2}} |D^\alpha \phi_N(t)|\, d^k t = o(N^{-1/2}).$$

Third step (the region $\varepsilon N^{1/2} < \|t\| \leqslant K_N N^{1/2}$). Since $|E \exp[it \cdot g(X_1)]| \leqslant 1 - x_2(\|t\|) < 1$, $t \neq 0$, we have

$$|E' \exp[it \cdot g'(X_1)]| = |E \exp[it \cdot g(X_1)]\, I(M(X_1) \leqslant N^{1/2})|(1 + o(N^{-3/2}))$$
$$\leqslant 1 - \chi_2(\|t\|) + O(N^{-3/2}). \tag{3.16}$$

Using the estimations of (3.15) for $t \in A_{N3}$ we have

$$|D^\alpha \phi_N(t)| \leqslant c\, N^{|\alpha|/2} \left(1 - \chi_2\left(\frac{\|t\|}{\sqrt{N}}\right) + o(N^{-3/2})\right)^{m-|\alpha|}$$

Hence there exists a sequence $K_N \uparrow \infty$ such that

$$|D^\alpha \phi_N(t) \leqslant c\, N^{-(k+1)/2} \eta_N \qquad \text{for} \quad \|t\| \leqslant N^{1/2} K_N,$$

thus proving

$$\int_{A_{N3}} |D^\alpha \phi_N(t)|\, d^k t = o(N^{-1/2}).$$

This together with the relation following (3.15) and (3.14) proves Lemma 3.10 and thereby Theorem 1.14.

Proof of Theorem 1.16. Here we don't need the third step of the proof of Theorem 1.14 and the proof simplifies for the expansion of the c.f. but is completely similar in the estimation step two.

Corollary 1.18 follows immediately by putting $g(X_1) = f_0(X_1) + N^{-1/2} f_1(X_1)$ in the proof of Theorem 1.14, while using the additional truncation

$$M(X_1) \triangleq E^{1/2}(\|f_2(X_1, X_2)\|^3 \,|X_1) + \|f_0(X_1)\| + N^{-1/2}\, \|f_1(X_1)\| \\ \leqslant N^{1/2}$$

which implies

$$\|g'(X_1)\| \leqslant N^{1/2} \qquad P'\text{-a.s.} \qquad \text{and} \qquad P(\|g(X_1)\| > N^{1/2}) = o(N^{-3/2}).$$

Furthermore,

$$|E' \exp[it \cdot (f_0(X_1) + N^{-1/2} f_1(X_1))]| \\ \leqslant 1 - \chi_2(\|t\|) + o(N^{-3/2}) + o(\|t\|\, N^{-1/2})$$

replaces (3.16) in the proof of Lemma 3.10(i) and a similar remark applies to the variance lower bound (2.6), thus proving Corollary 1.18.

Appendix

The following bound of moments is essential for the proof of Theorem 1.15:

LEMMA A.1. *For a given constant $C>1$ and $N \in \mathbb{N}$ let $H_{N,C}$ denote the class of symmetric kernels such that*

(i) $E(h(X_1, X_2) \mid X_2) = 0$, P-a.s.

(ii) $\|h(X_1, X_2)\| \leqslant C\,N$, P^2-*a.s.*,

(iii) $E\,\|h(X_1, X_2)\|^2 \leqslant C$, $E^{1/2}(\|h(X_1, X_2)\|^2 \mid X_2) \leqslant (C\,N)^{1/2}$, P-*a.s.*

Then

$$\sup\left\{E\left\|N^{-1}\sum_{\mu=1}^{N}\sum_{\nu=\mu+1}^{N} h_{\mu\nu}(X_\mu, X_\nu)\right\|^p : h_{\mu\nu} \in H_{N,C},\ N \geqslant 1\right\}$$
$$\leqslant C(p)(C + C^p) \qquad \textit{for every } p \geqslant 1.$$

Proof. Using the convexity of $p \to \log(E\,\|X\|^p)$, $p \geqslant 1$, we may assume w.l.g. that $p = 2m$, $m \in \mathbb{N}$. Using $(a+b)^m \leqslant 2^{m-1}(a^m + b^m)$ repeatedly it is sufficient to prove the assertion for dimension $k = 1$. Hence we have to estimate

$$S \triangleq \sum{}^* E[h_{i_1 i_2}(X_{i_1}, X_{j_1}) \cdots h_{i_{2m} j_{2m}}(X_{i_{2m}}, X_{j_{2m}})], \tag{A.2}$$

where $\sum^*$ denotes summation extending over all combinations of $2k$-tuples $1 \leqslant i_1, ..., i_{2k} \leqslant N$ and $1 \leqslant j_1, ..., j_{2k} \leqslant n$, such that $i_q < j_q$ for $q = 1, ..., 2k$. By virtue of (i) expectations of products in (A.2) containing a certain index only *once* have to vanish. Hence, we may assume w.l.g. that every index in these products occurs at least *twice*.

Let r denote the number of different indices among $i_1, ..., i_{2m}$, $j_1, ..., j_{2m}$. Assume w.l.g. that these are the indices 1, 2, ..., r and that the multiplicities of these indices satisfy

$$2 \leqslant n_1 \leqslant n_2 \leqslant \cdots \leqslant n_r. \tag{A.3}$$

We shall prove by induction on $r = 2, 3, ..., s$ that for all sets of indices satisfying (A.3) we have uniformly in $h_{\mu\nu} \in H_{C,N} \cup \Delta_{C,N}$, $\Delta_{C,N}$ as in (A.5)

$$\begin{aligned} P_{s,r} &= |E\,h_{i_1 j_1}(X_{i_1}, X_{j_1}) \cdots h_{i_s j_s}(X_{i_s}, X_{j_s})| \\ &\leqslant C^{r/2}(CN)^{s-r}, \end{aligned} \tag{A.4}$$

where $\mu = i_p = j_p$ is allowed, but then we require $h_{\mu\mu}(X_\mu, X_\mu) \in \Delta_{C,N}$, which means

$$E\,h_{\mu\mu}(X_\mu, X_\mu) \leqslant C, \qquad |h_{\mu\mu}(X_\mu, X_\mu)| \leqslant CN \qquad P\text{-a.s.} \tag{A.5}$$

The number of terms in (A.2) depending on r random variables X_j is of order $C(r)\,N^r$, hence assuming (A.4) we have

$$|S| \leqslant \sum_{r=2}^{2m} N^r\, C^{r/2}(CN)^{2m-r}$$
$$\leqslant c(m)(C+C^{2m})\,N^{2m}$$

thus proving Lemma A.1.

Proof of (A.4). When $r=2$, $m \geqslant 1$, $E|h_{12}(X_1, X_2)^{2m}| \leqslant (CN)^{2m-2}\, C^{2/2}$ by virtue of conditions (ii) and (iii).

Assume that (A.4) has been shown for $r=2,..., l-1$, $l \geqslant 3$. Taking conditional expectations with respect to X_1 (which has multiplicity $n_1 \geqslant 2$) yields, with

$$\mathscr{B}_1 \triangleq \sigma(X_j, j \neq 1)$$

for the expectation of all factors involving X_1,

$$\begin{aligned}|E(h_{1j_1}(X_1, X_{j_1}) \cdots h_{1j_{n_1}}(X_1, X_{j_{n_1}})|\, \mathscr{B}_1)\\ \leqslant (CN)^{n_1-2}\, C^{1/2}\, |\bar{h}_{j_1j_2}(X_{j_1}, X_{j_2})|,\end{aligned} \tag{A.6}$$

where

$$\bar{h}_{j_1j_2}(X_{j_1}, X_{j_2}) \triangleq C^{-1/2}\, E(h_{1j_1}(X_1, X_{j1})\, h_{1j_2}(X_1, X_{j2})\,|\,\mathscr{B}_1).$$

This kernel satisfies, by Hölder's inequality,

$$|\bar{h}_{j_1j_2}(X_1, X_2)| \leqslant C^{-1}(CN)^{2/2}, \qquad P^2-\text{a.s.}$$

as well as

$$E^{1/2}(\bar{h}_{j_1j_2}(X_1, X_2)\,|\,X_2) \leqslant (CN)^{1/2}.$$

Hence, $\bar{h}_{j_1j_2} \in H_{C,N}$ and since the remaining product in (A.4) for $r=l$ involves at most $s-2$ factors (plus one for $\bar{h}_{j_1j_2}$) there are exactly $l-1$ different X_j's. We have

$$P_{s,l} \leqslant (CN)^{n_1-2} C^{1/2}\, |E\, \bar{h}_{j_1j_2}(X_{j_1}, X_{j_2}) \prod_{p,q \neq 1} h_{pq}(X_p, X_q)|$$
$$\leqslant (CN)^{n_1-2}\, C^{1/2}\, C^{(l-1)/2}(CN)^{s-(n_1-1)-(l-1)}$$
$$\leqslant C^{l/2}(CN)^{s-l}$$

by induction hypothesis, which is applicable with $l-1$ provided (A.3) holds for $l-1$ indices. But the multiplicities n_j' of the remaining X_j's, $j>1$, after we carried out estimation (A.6) still satisfy $n_j' \geqslant 2$, $j > 1$. This follows by counting the number of times X_j occurs in (A.6) before and after estimation which shows

$$n_j - n_j' \leqslant n_1 - 2.$$

This together with $n_j \geqslant n_1$ immediately proves $n_j' \geqslant 2$.

REFERENCES

[1] BHATTACHARYA, R. N., AND RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

[2] BICKEL, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **4** 108–156.

[3] BICKEL, P. J., GÖTZE, F., AND VAN ZWET, W. R. (1984). The Edgeworth expansion for U-statistics of degree two. *Ann. Stat.* **14** (4).

[4] CALLAERT, H., AND JANSSEN, P. (1978). The Berry–Esseen theorem for U-statistics. *Ann. Statist.* **6** 417–421.

[5] CALLAERT, H., JANSSEN, P., AND VERAVERBEKE, N. (1980). An Edgeworth expansion for U-statistics. *Ann. Statist.* **8** 299–312.

[6] CARMICHAEL, W. P. (1981). *The Rate of Weak Convergence of a Vector of U-Statistics Generated by a Single Sample.* Ph.D. thesis, University of California, Berkeley.

[7] CHAN, Y.-K., AND WIERMAN, J. (1977). On the Berry–Esseen bound for U-statistics. *Ann. Probab.* **5** 136–139.

[8] HELMERS, R., AND JANSEN, P. (1982). *On the Berry–Esseen for Multivariate U-Statistics.* Mathematical Centre Report SW90, Amsterdam.

[9] HELMERS, R., AND VAN ZWET, W. R. (1982). The Berry–Esseen bound for U-statistics. In *Statistical Decision Theory and Related Topics, III,* Vol. 1 (S. S. Gupta and J. O. Berger, Eds.), pp. 497–512. Academic Press, New York.

[10] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[11] PFANZAGL, J. (WEFELMEYER, W.) (1985). Asymptotic expansions for general statistical models. In *Lecture Notes in Statistics,* Vol. 31. Springer-Verlag Heidelberg/New York.

[12] SWEETING, T. J. (1977). Speeds of convergence for the multidimensional central limit theorem. *Ann. Probab.* **5** 28–41.